Graph Spectral Clustering of Artefacts In Radio Interferometric Images

Matthieu Simeoni *,†, Paul Hurley ‡

* IBM Zurich Research Laboratory, Rüschlikon

† École Polytechnique Fédérale de Lausanne (EPFL), Lausanne

‡ Western Sydney University, Australia

IBM EPFL WESTERN SYDNEY UNIVERSITY

IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2019)

Abstract — Raw radio interferometric images are severely **blurred** by strong convolution artefacts. In this work, we propose an **unsupervised learning** method for clustering such artefacts. We use spectral clustering on a **kinship** graph: each pixel interpreted as a node, linked by edges weighted according to their spatial correlation. Actual celestial sources are then **identified** within each clusters.

1. Sky Covariance Function

Assume a **point source field** *S*, with *randomly* fluctuating amplitudes

$$S(\boldsymbol{r}) = \sum_{q=1}^{Q} \xi_q\, \delta(\boldsymbol{r}-\boldsymbol{r}_q), \quad \forall \boldsymbol{r} \in \mathbb{S}^2$$

The radio telescope **measurements** are given by

$$\boldsymbol{y} := \sum_{q=1}^{Q} \xi_q \begin{bmatrix} e^{-\frac{2\pi j}{\lambda}\langle \boldsymbol{r}_q, \boldsymbol{p}_1\rangle} \\ \vdots \\ e^{-\frac{2\pi j}{\lambda}\langle \boldsymbol{r}_q, \boldsymbol{p}_L\rangle} \end{bmatrix} = \sum_{q=1}^{Q} \xi_q \boldsymbol{\varphi}(\boldsymbol{r}_q).$$

Evidence of the *steering vector* in the data reveal *likely positions* of the underlying sources:

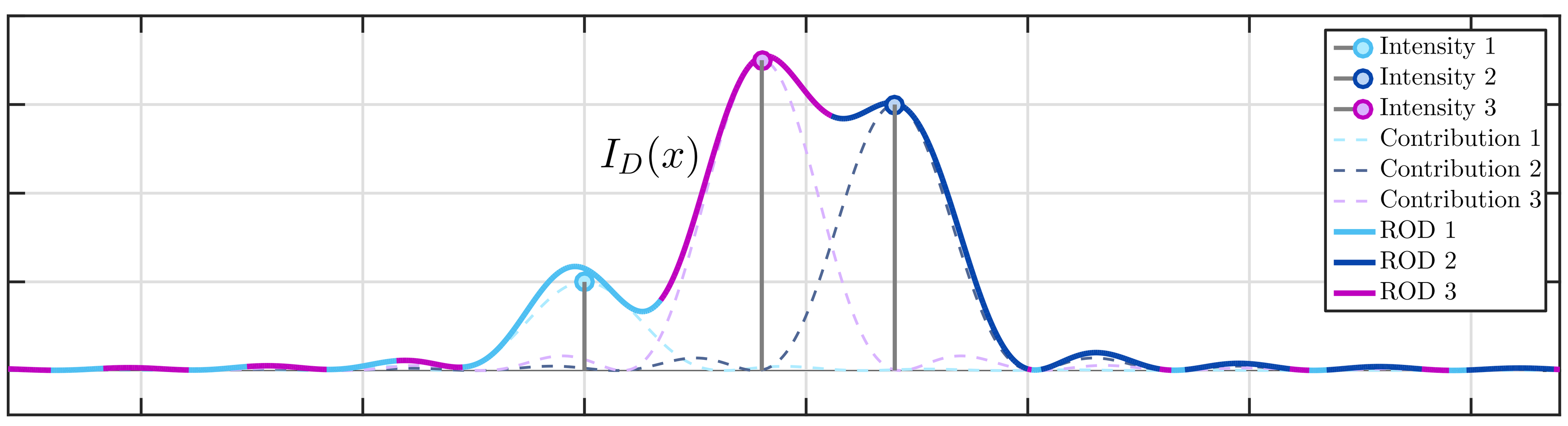
$$\hat{S}(\boldsymbol{r}) = \langle \boldsymbol{y}, \boldsymbol{\varphi}(\boldsymbol{r})\rangle = \boldsymbol{\varphi}(\boldsymbol{r})^H \boldsymbol{y}, \quad \boldsymbol{r} \in \Theta.$$

Dirty field is characterized by its **covariance kernel**

$$\hat{\kappa}(\boldsymbol{r}, \boldsymbol{s}) := \mathbb{E}\left[\hat{S}(\boldsymbol{r})\hat{S}^*(\boldsymbol{s})\right] = \boldsymbol{\varphi}(\boldsymbol{r})^H \Sigma \boldsymbol{\varphi}(\boldsymbol{s}).$$

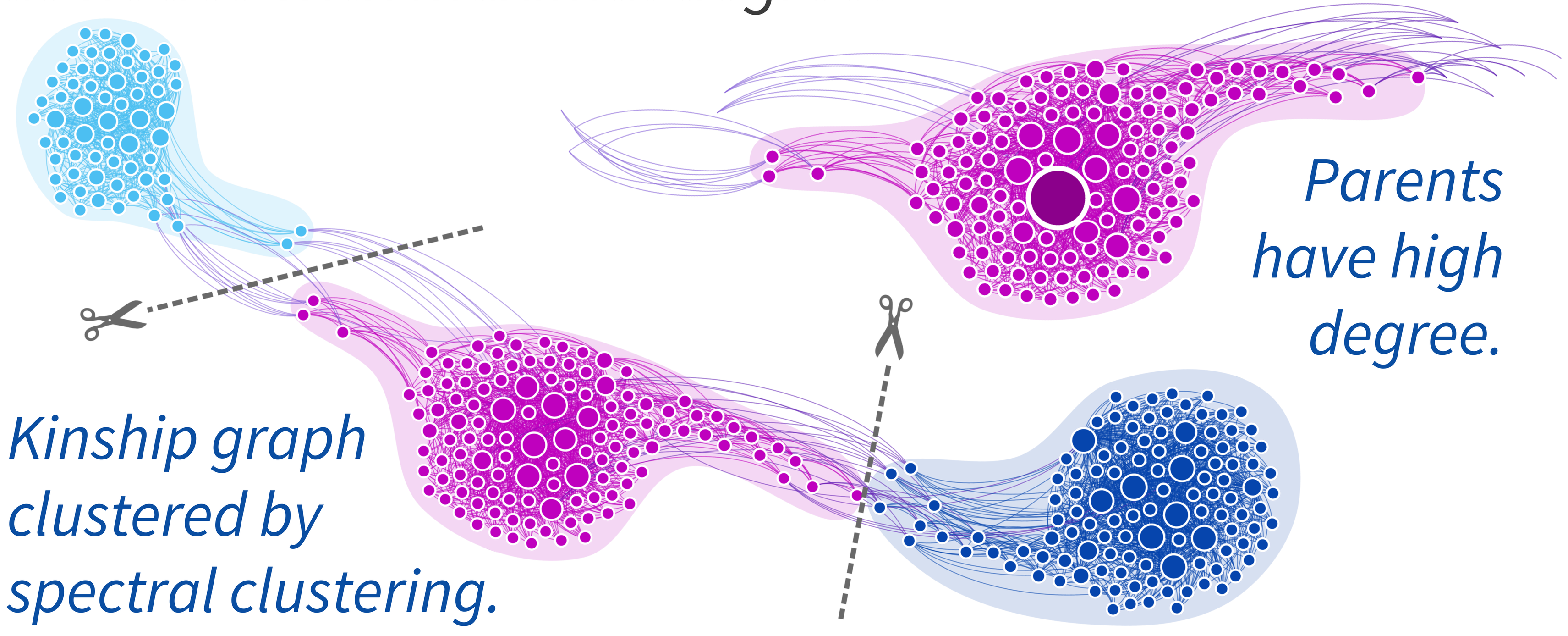
2. The Parenting Problem

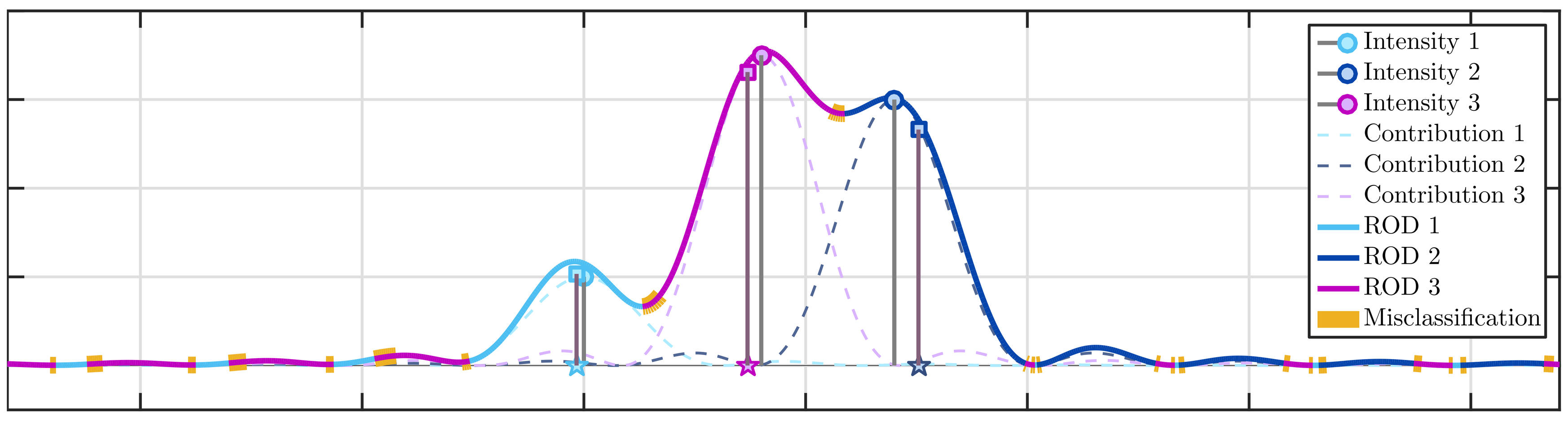
Radio astronomers only look at **intensity field** (variance). It is polluted by strong **convolution artefacts:** each source leaks its energy in the field. In practice, we observe **regions of dominance.**



3. Kinship Graph & Spectral Clustering

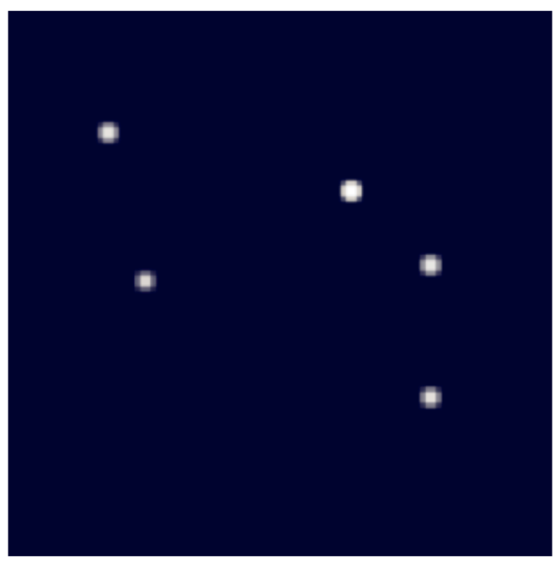
We define a **kinship graph**, linking each sky direction (node) according to their **degree of kinship** (correlation). The goal is to *cluster* this kinship network to recover the region of dominances. We use *spectral clustering*. **Parent sources** are identified as nodes with *maximal degree*.



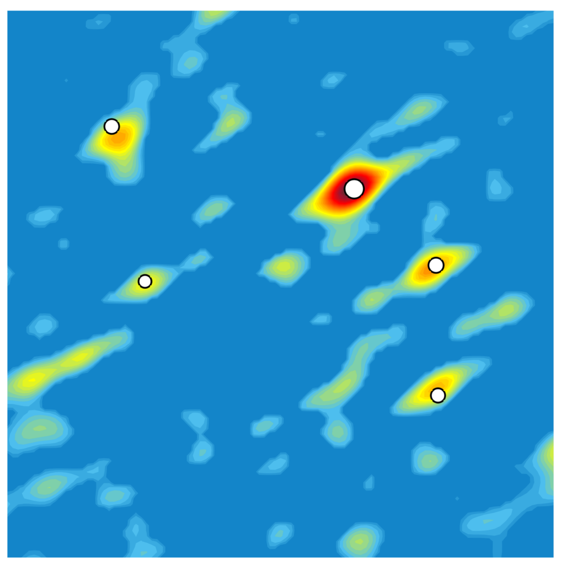


Region of dominances recovered by spectral clustering (90% accuracy)

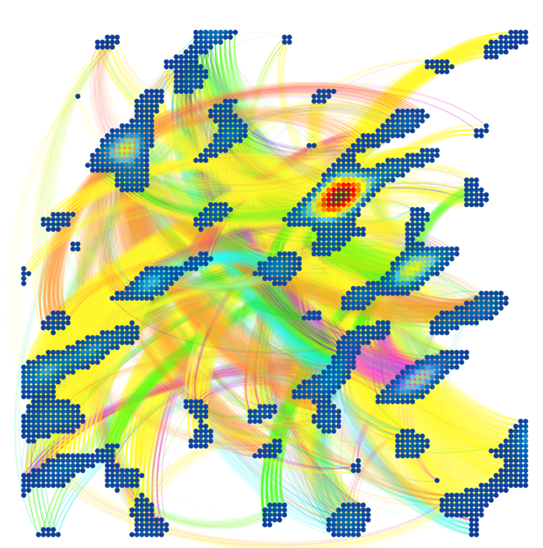
4. Experimental Results



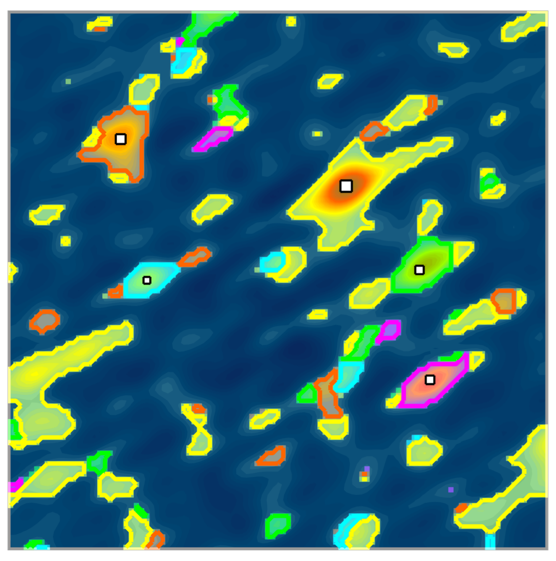
True sky.



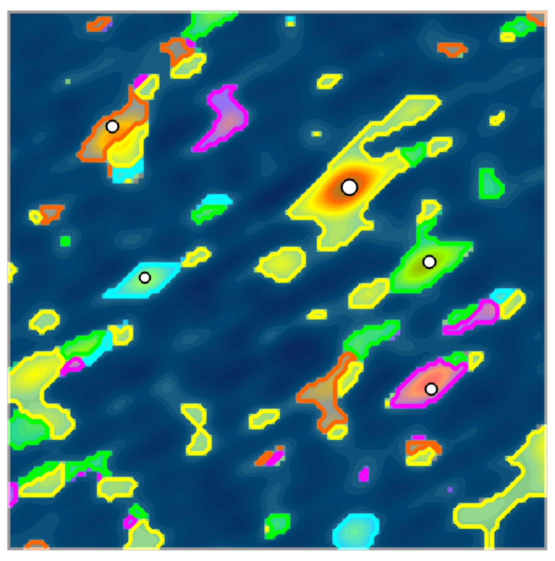
Dirty image.



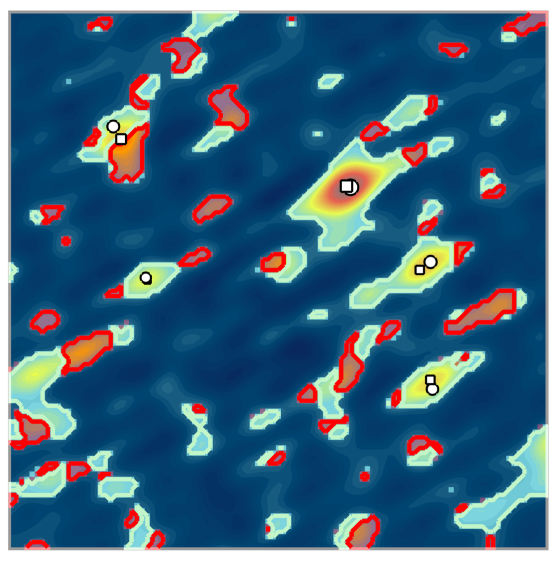
Kinship graph.



True ROD.



Recovered ROD.



Accuracy (86.7 %)